

Philosophiques

philosophiques

Denis Dennett, *La stratégie de l'interprète : le sens commun et l'univers quotidien*, Gallimard, NRF Essais, Traduction française de P. Engel, 493 pages.

Luc Faucher

Volume 19, numéro 1, printemps 1992

URI : <https://id.erudit.org/iderudit/027186ar>

DOI : <https://doi.org/10.7202/027186ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Société de philosophie du Québec

ISSN

0316-2923 (imprimé)

1492-1391 (numérique)

[Découvrir la revue](#)

Citer ce compte rendu

Faucher, L. (1992). Compte rendu de [Denis Dennett, *La stratégie de l'interprète : le sens commun et l'univers quotidien*, Gallimard, NRF Essais, Traduction française de P. Engel, 493 pages.] *Philosophiques*, 19(1), 162–168.
<https://doi.org/10.7202/027186ar>

Daniel DENNETT, *La stratégie de l'interprète: le sens commun et l'univers quotidien*, Gallimard, NRF Essais, Traduction française de P. Engel, 493 pages.

par Luc Faucher

La traduction française du texte de Dennett est bienvenue. Elle permettra au lecteur francophone d'avoir un aperçu des principales polémiques contemporaines en philosophie de l'esprit. Rappelons pour ceux qui ne le saurait pas déjà que l'œuvre de D. C. Dennett tourne principalement autour de deux questions: l'intentionnalité et la conscience. Dans *La stratégie de l'interprète*, Dennett choisit de ne

traiter qu'un seul de ces deux thèmes: l'intentionnalité. Un autre livre, *Consciousness Explained* (1991), est consacré à la question de la conscience. Il est significatif que Dennett ait choisi de traiter les deux sujets indépendamment l'un de l'autre, alors qu'au même moment, des philosophes comme Searle avouent avoir commis une erreur en négligeant la relation entre conscience et intentionnalité (« Consciousness, Explanatory Inversion and Cognitive Science », *BBS*, 1991). C'est que Dennett préfère d'abord construire une théorie de l'intentionnalité ou du contenu en laissant de côté la question de la conscience, puis construire une théorie de la conscience sur cette théorie de l'intentionnalité. Dennett considère en effet que dans la plupart des cas, l'intentionnalité n'est pas transparente, c'est-à-dire que la plupart des états intentionnels d'un système sont « à propos » de quelque chose sans que le système en soit conscient.

Cette traduction de *The Intentional Stance* regroupe en trois sections distinctes des essais pour une grande part déjà publiés, mais complétés dans ce cas de réflexions nouvelles. Les essais de la première section (chap. I à IV) sont consacrés au développement de la stratégie intentionnelle. La seconde section (chap. V) avance un certain nombre d'arguments destinés à montrer les difficultés d'une version naïve du langage de la pensée. La troisième section (chap. VI à X) inclut des essais où Dennett met en évidence la parenté entre le modèle adaptatif en biologie et la stratégie intentionnelle. Le lecteur qui voudrait avoir un aperçu rapide des positions récentes sur l'intentionnalité ainsi que la place qu'occupe Dennett dans ce débat aurait avantage à consulter d'abord le dernier chapitre de l'ouvrage où Dennett établit une cartographie précise permettant de situer les principaux philosophes de l'esprit selon leur position sur le sujet.

La préoccupation constante qui traverse ces sections est la question du type de réalité des états attribués à l'aide des concepts de la psychologie populaire. Rompant avec une position instrumentaliste défendue dans *Brainstorms*, Dennett suggérera que la stratégie intentionnelle s'appuie sur une forme de réalisme tempéré. Selon Dennett, les attitudes propositionnelles sont des *abstracta* (entités théoriques liées au calcul de la théorie) qui permettent de saisir des régularités abstraites de l'environnement et qui ne correspondent pas nécessairement à des régularités physiques. Les régularités mises au jour par la psychologie populaire n'apparaissent qu'à ceux qui adoptent la posture (le point de vue) intentionnelle. Une fois adoptée, la posture met au jour certains schémas d'action réels, mais inaccessibles par le biais d'autres stratégies prédictives. L'irréductibilité des schémas n'est pas une preuve de leur inexistence. Le fait que Dennett veuille soutenir leur existence le force à soutenir un moyen terme entre le réalisme fort et l'éliminativisme. La position réaliste forte, dont Fodor est un tenant, suggère que les désirs et les croyances ne sont réels que si les régularités discernables à l'aide de la psychologie populaire sont isomorphes aux régularités au niveau des structures du cerveau.

L'éliminativisme, défendu entre autres par Paul Churchland, repose pour sa part sur l'espoir que les neurosciences découvriront des régularités si clairement supérieures aux régularités de la psychologie populaire que nous serons forcés d'abandonner cette dernière. Contre la position de Fodor, Dennett tentera de montrer que la correspondance un à un entre les états intentionnels isolés à l'aide de la psychologie populaire et les états du cerveau est au mieux une hypothèse empirique et au pire une utopie. Dennett reprochera à la théorie de Churchland d'être affectée par une myopie extensionnaliste, c'est-à-dire de perdre l'idée de simplification et d'abstraction, qui rendait si attrayante la psychologie populaire, au profit d'une description physique extensive. Les raisons qui nous permettent de ne pas espérer l'existence d'états discrets qui correspondent aux états attribués à l'aide de la stratégie de Dennett me semblent assez convaincantes (voir plus bas). Il reste cependant à prouver que cette stratégie est pratiquement indispensable. Passons maintenant au contenu proprement dit des essais.

« Les vrais croyants » constitue l'introduction à la théorie de l'esprit de Dennett (elle remplace « *Intentional Systems* » (1971)). L'approche de Dennett est fondée sur la distinction entre les stratégies que l'on peut adopter pour expliquer le comportement d'artefacts ou de créatures. Une stratégie intentionnelle consiste à considérer un objet dont on veut prédire le comportement comme pouvant être décrit à l'aide d'états intentionnels et comme étant optimalement rationnel. Un objet sera un système intentionnel si on peut prédire avec succès son comportement à l'aide de cette stratégie. L'intérêt de la posture intentionnelle est qu'elle nous permet de faire des prédictions qui nous étaient impossibles ou difficiles d'un point de vue purement physique ou fonctionnel (qui sont en fait les deux autres stratégies intéressantes de prédiction du comportement, selon Dennett). Cette stratégie nous est naturelle, mais comme le notait Quine récemment, le fait qu'elle nous soit naturelle n'assure en rien sa vérité. En fait, il n'est pas nécessaire de postuler que les catégories qu'elle utilise pour réaliser ses prédictions soient des noms d'espèces naturelles. La stratégie ne vise pas tant à décrire ce qui se produit à l'intérieur du cerveau, qu'à fournir la prédiction la plus exacte possible du comportement d'un système à partir des régularités observables.

La stratégie intentionnelle est celle que nous utilisons couramment dans nos transactions avec autrui, lorsque nous prédisons son comportement à l'aide de croyances et de désirs, et que nous déduisons son comportement en assumant sa rationalité. Cette stratégie a toutefois des ratés. Ce sont ces cas qui nous forcent à concevoir le système auquel on a affaire comme un agent épistémique non idéal, c'est-à-dire dans un premier temps, à ne lui attribuer que les croyances qui sont pertinentes relativement à ses intérêts et les désirs des choses qu'il croit bonnes, et dans un second temps, à réviser sa rationalité à la baisse (il n'est pas nécessaire qu'un sys-

tème puisse tirer les implications de toutes ces croyances, mais seulement les implications pertinentes à sa survie). Cette version affaiblie marque une certaine distance prise par la stratégie de Dennett avec la psychologie populaire. La relation proposée dans « Trois sortes de psychologie intentionnelle » entre la psychologie populaire et sa stratégie en est une de réduction. Cette réduction a pour but d'assurer une description systématique et entièrement générale du domaine un peu flou du mental qui émane de la psychologie populaire. La théorie des systèmes intentionnels permettrait de décrire les compétences de certaines créatures en termes abstraits, alors qu'une théorie cognitive subpersonnelle serait chargée d'expliquer comment des systèmes particuliers peuvent exhiber ces compétences. La théorie cognitive serait aussi chargée d'expliquer pourquoi, dans certains cas, le système ne se comporte pas comme un système idéal (par exemple, expliquer en termes d'heuristiques sous-optimales les problèmes d'irrationalité mis au jour par Tversky et *al.*). La théorie cognitive subpersonnelle serait aussi chargée d'expliquer comment une machine syntaxique comme le cerveau peut donner l'impression d'être une machine sémantique. Ce dernier type de psychologie ne sera pas solipsiste, puisqu'une grande part de son explication va porter sur la façon dont les « systèmes qui *semblent* distinguer des significations [distinguent] en fait des choses (des *tokens* de types manifestement très disjoints) qui co-varient de manière fiable avec des significations » (p. 86). Cette dernière affirmation a deux implications: d'abord, comme nous l'avons suggéré, la psychologie devra chercher à savoir comment l'évolution s'y est prise pour « calibrer » les états syntaxiques de l'organisme avec le monde; puis, elle devra chercher comment la causalité mise au jour au niveau sémantique ou intentionnel, est expliquée par le fonctionnement du niveau syntaxique.

Dans « Styles de représentation mentale » Dennett précise le type de représentation mentale auquel devra faire référence cette psychologie cognitive subpersonnelle. Dennett propose d'abord un ensemble de distinctions importantes à propos des types de représentation ou d'information. Il existe selon lui trois types de représentation mentale: les représentations explicites, implicites et tacites. Une représentation est explicite si à un endroit d'un système il existe un objet « [...] physiquement structuré, une formule, ou une suite de formules, ou une inscription de *tokens* de certains membres d'un système (ou « langage ») d'éléments pour lesquels il y a une sémantique ou interprétation, et une « provision » [...] pour lire ou pour déchiffrer la formule » (p. 283). Une représentation est implicite si elle peut être dérivée à partir des représentations explicites. Le troisième type de représentation, de l'aveu même de Dennett, n'a jamais été clairement défini. Une règle tacite, pour prendre cet exemple, est une règle qui est « honorée » par un système, mais qui n'est ni représentée explicitement, ni dérivée de représentations de règles explicites. Les systèmes connexionnistes sont, selon

Dennett, des systèmes qui peuvent être décrits comme ayant des représentations tacites. L'erreur de la plupart des cognitivistes (dont Chomsky et Fodor) est d'avoir voulu utiliser des règles explicites pour représenter le savoir des systèmes qu'ils voulaient décrire. Quoique l'enthousiasme de Dennett pour le connexionnisme soit tempéré, il croit que les systèmes de ce type risquent d'être des concurrents sérieux à l'approche cognitive. En effet, nulle part dans un système connexionniste ne trouve-t-on de règles explicitement représentées. Ces règles sont tacitement représentées dans « la structure dispositionnelle émergente du réseau » (p. 301). On ne peut donc pas espérer *a priori* que les catégories de la psychologie populaire, et les règles et représentations dérivées à l'aide de celle-ci, se retrouveront dans la structure des systèmes que l'on peut décrire en ces termes. La psychologie cognitive subpersonnelle proposée plus haut utilisera donc des représentations dont le type ne sera pas computationnel.

Le second texte de la troisième section, « Les systèmes intentionnels en éthologie cognitive », permettra au lecteur de prolonger sa réflexion sur le statut scientifique de la stratégie de Dennett. Ce texte s'ouvre sur une proposition de Dennett aux éthologistes — qui ont rencontré dans leurs pratiques les limites du béhaviorisme — d'utiliser sa théorie des systèmes intentionnels pour générer des hypothèses empiriques sur le comportement des animaux. Dennett entend aussi montrer que les analyses adaptationnistes sont du même genre que les analyses intentionnelles et, par le fait même, défendre l'approche adaptationniste. La première proposition va de soi, puisque la tâche assignée à la méthode de Dennett est largement semblable à celle des éthologistes: interpréter le comportement d'un système de l'extérieur à l'aide d'un vocabulaire intentionnel permettant d'obtenir des pouvoirs prédictifs que ne permet pas le béhaviorisme. Cette proposition ne me semble pas faire problème; la seconde proposition est, elle, beaucoup plus discutée en philosophie de la biologie (Gould et Lewontin ont tenté de montrer que l'adaptationnisme adopte une vue trop étroite des processus à l'oeuvre dans la sélection naturelle). La façon qu'a Dennett de sauver sa stratégie des critiques de ce type est astucieuse: elle consiste à assimiler leur type de critique à celui de Skinner envers le mentalisme et de les condamner par association. Gould et Lewontin auraient voulu éliminer l'explication adaptationniste, pour les mêmes raisons que Skinner avait éliminé l'explication mentaliste, c'est-à-dire parce que les phénomènes qu'elle postule risquent de n'être que des fabulations infalsifiables. Selon Dennett, l'explication adaptationniste et la stratégie de l'interprète ne sont pas des théories au sens fort, qui devraient « enluminer la réalité », comme l'écrit Engel, mais des points de vue qui nous permettent de décrire des compétences ou des histoires évolutives possibles qui donneraient à la science des pistes pour engager la recherche. Ces stratégies n'ont pas à périr en même temps que leurs hypothèses, puisqu'el-

les ne prétendent pas décrire les mécanismes ou les processus effectifs à l'œuvre dans un système ou dans le processus de sélection naturelle. Il faudrait toutefois prouver l'indispensabilité pratique de ces stratégies, ce que je ne suis pas sûr que Dennett ait réussi à faire.

« Évolution, erreur et intentionalité » s'attaque à une position qui travaille la philosophie de l'esprit contemporain: l'attitude à adopter quant à l'interprétation des artefacts et des humains. *Grosso modo*, la question est de savoir si une machine ou un texte ont une intentionalité du même type que la nôtre, ou s'ils possèdent une intentionalité qui dérive de la nôtre. Les Searle, Fodor, Drestke et Kripke — entre autres — ont opté pour la seconde alternative. Le corollaire de leur position est que nous possédons une intentionalité non dérivée. Dennett propose l'expérience de pensée suivante: on dépose votre corps en hibernation dans un robot conçu pour assurer votre survie pour les siècles à venir. La meilleure stratégie pour permettre la survie, est de concevoir un robot qui peut se déplacer, qui peut tirer des plans à partir de ces buts et qui pourrait, si le besoin s'en fait sentir, coopérer avec d'autres robots. S'appuyant sur la similarité de cette situation avec celle des gènes égoïstes de Dawkins (selon ce dernier, les gènes se servent des corps comme des véhicules permettant leur perpétuation), il suggère que nous sommes ce type de robot et que par conséquent notre intentionalité dérive de celle de Mère Nature. Cela a deux conséquences majeures: d'abord la signification de nos états mentaux est indéterminée et ensuite, l'intentionnalité originale caractérise l'intentionnalité de Mère Nature (terme désignant le processus non intentionnel de la sélection naturelle). On a reproché à Dennett (Rudder Baker, « Recent Work in Philosophy of Mind », *Philosophical Books*, 1989) de commettre un cercle en faisant dépendre notre intentionalité de celle de Mère Nature. Ce reproche n'est possible que si on oublie que notre intentionalité dépend de celle de Mère Nature, mais dans le sens où nous ne pouvons faire d'attributions si nous ne postulons pas une certaine forme d'optimalité dans la nature assurant la rationalité du système visé par l'attribution. Notre intentionalité ne dérive pas de celle de Mère Nature en un sens causal, mais seulement logique, ce qui ne nous force pas à exiger que ses intentions soient déterminées.

Au dernier chapitre de *La stratégie de l'interprète*, Dennett tente de placer les différentes positions actuelles en philosophie de l'esprit à l'intérieur d'un *continuum* dont les extrémités seraient occupées par les positions respectives de Sellars et Quine sur la question des attitudes propositionnelles. On doit d'abord à ces philosophes la formulation d'une paire de thèses antithétiques autour de la question de l'intentionnalité: ou bien la thèse de Brentano démontre le besoin d'une science de l'intention ou bien elle démontre l'inanité du langage intentionnel. Quine, qui opta pour la seconde

position dans *Word and Object*, proposa toutefois une avenue médiane: le langage intentionnel peut être conservé pour des raisons pragmatiques, mais sa valeur pour une science extensionnelle est nulle. En effet, le langage mentaliste est affecté d'une indétermination qui fait que l'on ne peut espérer que corresponde aux catégories de ce langage des processus physiques discrets. Une seconde distinction, dont Quine peut encore revendiquer la paternité, doit être faite si on accepte la position mitoyenne. Si le langage intentionnel n'a qu'une fonction interprétative, on peut encore se demander de quel type est l'interprétation, normative (comme dans le principe de charité) ou projective (comme dans l'empathie ou la théorie de la simulation). Une fois ces deux distinctions établies, les positions des Fodor, Dennett, Stich, Churchland et Davidson peuvent être placées à divers points du *continuum*.

La position de Dennett s'inscrit donc en continuité avec celle de Quine. Comme ce dernier, il constate que le langage intentionnel est indéterminé et que le langage de la science doit être purement extensionnel. Ces constatations ne conduisent cependant pas Dennett à désirer la disparition du langage intentionnel. Elles le conduisent par contre à reconnaître l'improbabilité que les catégories décrites à l'aide de la stratégie intentionnelle aient des correspondants physiques déterminés. Ce serait se méprendre sur la portée de la stratégie si on prétendait qu'elle constitue une description des états internes d'un système. Cette stratégie peut-être mieux comprise si on la voit comme une façon de caractériser le système vis-à-vis de son environnement. Vu sous cet angle, la stratégie intentionnelle ne risque pas de disparaître avec l'avènement du connexionnisme (comme semble le craindre Fodor et l'espérer les Churchland), elle peut continuer à être considérée comme une description du système qui l'inclut dans son environnement.

Les autres essais, que je n'ai pas eu le temps de discuter, sont « Partir du bon pied » (profession de foi optimiste de Dennett envers une philosophie naturalisée), « Donner un sens à ce que nous faisons » (qui est une réponse à une critique de Stich qui permet à Dennett de clarifier sa position vis-à-vis de sa thèse de la rationalité optimale), « Au-delà de la croyance » (où Dennett propose un concept « d'attitudes notionnelles » qui aurait pour lui la clarté que n'ont pas les concept « d'attitudes propositionnelles » et « d'attitudes phrastiques ») et « Pensée rapide » (qui montre que l'attrait de l'exemple de la chambre chinoise de Searle est un attrait pour la conclusion, non pour l'argument de Searle). Chacun de ces essais propose des perspectives intéressantes qui valent elles aussi d'être considérées.

*Département de philosophie
Université du Québec à Montréal*